FREQUENCY CURVES OF CLIMATIC PHENOMENA.[1]

By Howard Ross Tolley, Scientific Assistant.

[Dated: U. S. Office of Farm Management, Washington, Oct. 24, 1916.]

A person who wishes to attempt farming in a new region, or even one who farms intelligently in a region with which he is acquainted, must have more than a haphazard knowledge of the variability of the climate of the region. He should know how often it will be too wet for one crop or too dry for another; how often the summer will be too cool or too hot; how often the growing season will be too short or the winter too cold. In short, a knowledge of the frequency of unfavorable occurrences of climatic phenomena is necessary for successful agriculture, and inasmuch as this knowledge enters into the determination of what crops are or are not suitable for any given locality, the investigation of the variability of climatic phenomena is a matter of great importance in farm management.

All farmers know that their business is liable to loss on account of unfavorable weather, and most of them have a fair empirical knowledge of the amount of risk to which their crops are subject. To the man with only a small capital, the loss of a single staple crop will often be disastrous, and the unsuccessful farmer can often trace his failure to injurious weather. Of course, this hazard can not be controlled, but a knowledge of the ways in which climatic phenomena occur in general, and an investigation of the records for a particular locality will render it possible at least to reduce the risk of such losses. Maps showing average conditions are at present available, but the average alone is of little practical significance unless supplemented by data showing the relation between the average and the actual occurrences. Maps or other methods presenting data in concise form that would show the variations of the different phenomena from their averages would greatly increase the value of the maps now available.[2]

Where the weather has been under observation at or near a place for a sufficient length of time to enable one to determine the average conditions accurately,[3] an examination of the record will reveal something as to the frequency with which a phenomenon may be expected to occur in different ways. The records of all the phenomena at all the stations should be handled in the same way, a general method being employed in all cases, so that the results may be comparable. In this, as in all other statistical work, the use of frequency curves offers the most systematic method of examining the variations present in a series of observations. This method is necessarily mathematical, and in regard to the use of mathematical processes in this, or any other kind of investigation, it can only be said that—

They are the abbreviators of long and tedious operations, and it would be perfectly possible, with sufficient time and industry, to do without their use. * * * When both the ordinary and mathematical results are derived from the same hypothesis, the latter must be the more correct; and in those numerous cases in which the difficulty lies in reducing the original circumstances to a mathematical form, there is nothing to show that we are less liable to error in deducing a common-sense result from principles too indefinite for calculation than we should be in attempting to define more closely and apply numerical reasoning.[4]

The frequency distributions, or the resulting frequency polygons of the records of most climatic phenomena, show that there is a tendency for the number of occurrences to become greater as the middle point is approached from either end, and the value that occurs oftener than any other (the mode) will generally be found to be somewhere near the average of all the observations (the mean), and the value dividing the occurrences into halves (the median) will also be near the mean.[5] The usefulness of the average is largely dependent on the assumption that these three values are very nearly, if not exactly, coincident. If the number of observations is great, i. e., if the record is a very long one, the distribution of the different frequencies will be more regular than if there are only a few observations. If it were possible to increase the number of observations indefinitely, all the irregularities would disappear and the frequency polygon would very closely approach a smooth curve, rising gradually from the base line at a point beyond that representing the lowest observed value of the variable, reaching its highest point at the value that occurred most frequently, then gradually falling away to the base line again above the highest observed value. If, in this limiting case, the polygon, or curve representing it, is symmetrical about the ordinate of the mean as an axis, the median and mode coincide with the mean, i. e., the average of all the observations is the value that occurs most frequently, and there are the same number above the average as below it. However, if the curve is not symmetrical, the mean and the mode will not coincide and the distance between them will depend upon the amount of deviation from symmetry, or skewness, present in the distribution.

In practical work it is not possible to increase the number of observations at will and the data must be used as they are found, but it may be possible to determine the ideal curve which a given frequency distribution approaches. This curve will give a frequency corresponding to any value of the variable, and if it is known, the investigator is no longer restricted to the use of the arbitrary groups of the observations themselves. From the examination of a limited number of observations, then, it becomes possible to obtain a reasonable estimate of the series of frequencies that would result from an unlimited series of observations. That is, an examination of the record of any climatic phenomenon might enable one to form an idea of the distribution to be expected.

If it can be satisfactorily demonstrated that in the long run the average value will occur more frequently than any other, and that deviations above and below the average are equally likely to occur, it is generally safe to assume that the ideal form of the distribution is that shown by the normal frequency curve. The characteristics of any distribution which can be represented by this curve can all be expressed by a single number, the standard deviation, and in order to determine the curve that will represent the ideal form of any series of data that exhibits the properties indicated above, it is only necessary to know the position of the mean and the size of the standard deviation. From this it is possible to find the portion of the total number of occurrences which will have a value greater or less than any selected amount, or the portion that will be between any two amounts; or, what is the same thing, the probability that any observed value will be greater or less than a given value, or that it will lie between any two values. Tables[6] have been

[1] The writer wishes to record his obligation to Prof. W. J. Spillman, Chief U. S. Office of Farm Management, at whose suggestion and under whose direction this study has been undertaken.
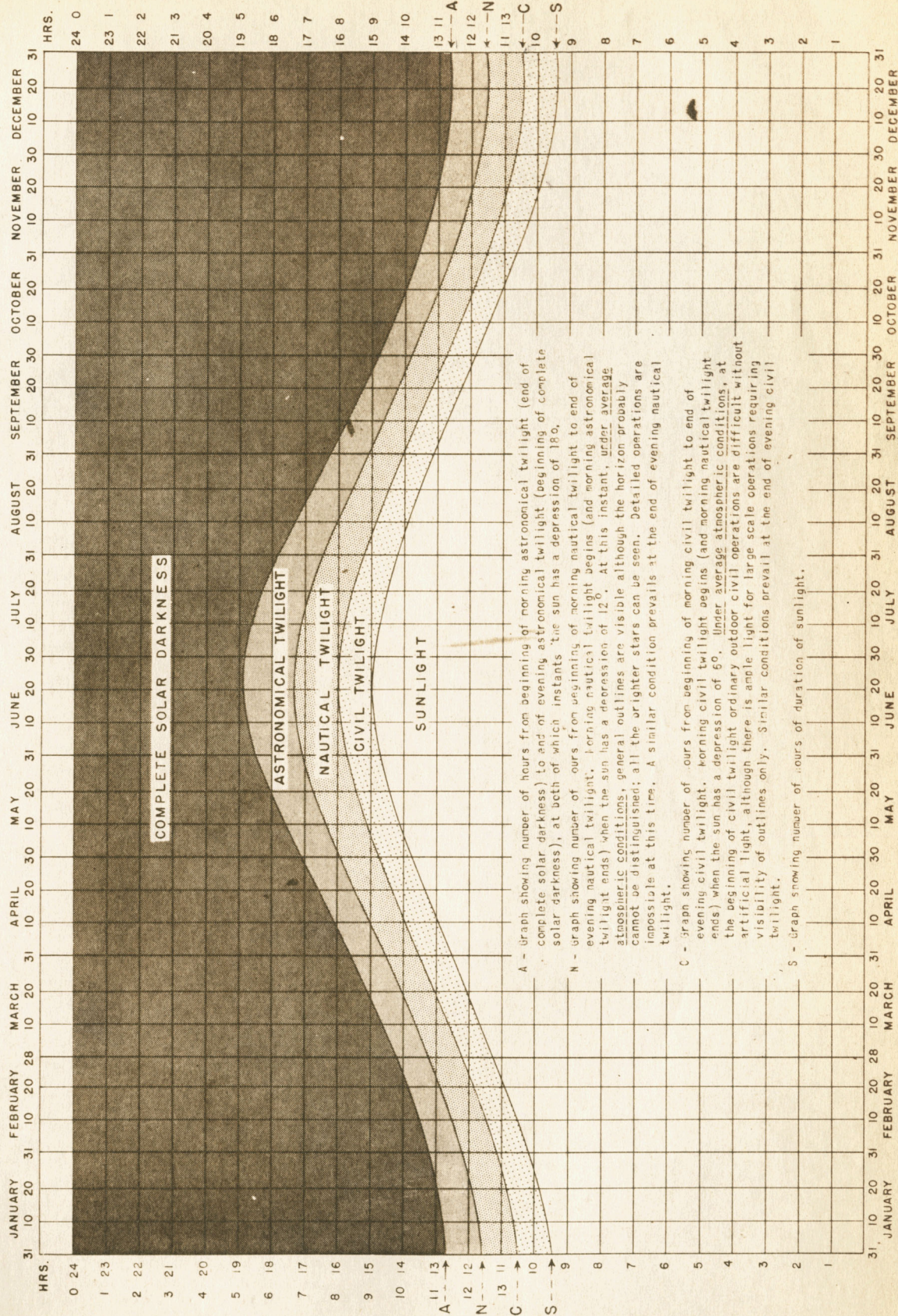
[2] In this connection, see Reed, W. G. The probable growing season. Monthly Weather Review, Sept., 1916, 44: 509-512, and charts.

[3] The number of observations required to deduce an accurate average depend somewhat upon the range of their values, but in any case, little weight should be given to an average obtained from less than twenty observations.

[4] De Morgan, Augustus, 1806-1871.

[5] For a more detailed discussion with examples, see C. F. Marvin, Elementary notes on least squares * * * for meteorology and agriculture. Monthly Weather Review, Oct., 1916, 44: 551-569.

[6] Davenport, C. B. Statistical methods. ed. 3, New York, 1914. p. 119.
Pearson, K. Tables for statisticians and biometricians. Cambridge, 1914. p. 2.

WASHINGTON, D.C.
LAT. 38°50'N

COMPLETE SOLAR DARKNESS

ASTRONOMICAL TWILIGHT

NAUTICAL TWILIGHT

CIVIL TWILIGHT

SUNLIGHT

A – Graph showing number of hours from beginning of morning astronomical twilight (end of complete solar darkness) to end of evening astronomical twilight (beginning of complete solar darkness), at both of which instants the sun has a depression of 18°.

N – Graph showing number of hours from beginning of morning nautical twilight to end of evening nautical twilight. Morning nautical twilight begins (and morning astronomical twilight ends) when the sun has a depression of 12°. At this instant, under average atmospheric conditions, general outlines are visible although the horizon probably cannot be distinguished; all the brighter stars can be seen. Detailed operations are impossible at this time. A similar condition prevails at the end of evening nautical twilight.

C – Graph showing number of hours from beginning of morning civil twilight to end of evening civil twilight. Morning civil twilight begins (and morning nautical twilight ends) when the sun has a depression of 6°. Under average atmospheric conditions, at the beginning of civil twilight ordinary outdoor civil operations are difficult without artificial light, although there is ample light for large scale operations requiring visibility of outlines only. Similar conditions prevail at the end of evening civil twilight.

S – Graph showing number of hours of duration of sunlight.

WRC-SR-1843

constructed from which it is possible to find these quantities directly when the mean and standard deviation have been determined. It is evident that the use of this property of the normal curve will give much more satisfactory results than a simple count of the number of such occurrences during a given interval of time, for the shortness of the records of all climatic phenomena makes it impossible to assume that the percentage of cases which exceed a given value, say, is the same as the percentage that would be observed in the long run.

As an example, take the case of the last killing frost in Spring, the distribution of which follows the normal frequency curve, and suppose it is desired to find the ferquency with which frost should be expected after any given date at a particular station.[7] The longest available frost record in the United States covers a period of only 59 years, and it is not reasonable to assume that the frequency distribution of a record of this length is regular enough to warrant one in drawing any very close deductions from a mere count of the number of times the last frost occurred on the different dates. However, by using the frequency curve of the distribution in the manner above described, one may avoid this assumption. The form of the curve depends on the standard deviation which, in turn, is determined by the way all the observations are arranged, and consequently the difficulties due to the irregularities which are always present in the distribution of a small number of observations disappear. The most reliable method of finding the date after which a certain portion, say 1/20, of the last frosts should be expected to occur, would be to find the date corresponding to the point whose ordinate cuts off 1/20 of the area of the frequency curve for the distribution. An examination of the tables shows[8] that the ratio of the departure from the mean to the standard deviation, x/σ, of such a point is 1.645. That is, the date after which the last killing frost will occur in 1/20 of the years is 1.645σ days after the observed average date.

This method of procedure is comparatively simple, and where the distribution of a phenomenon is truly normal, it is possible to obtain very reliable statements of the risk from unfavorable conditions. However, the frequency distributions of most climatic phenomena are not normal, and in such cases the procedure is not nearly so simple and clear cut. A convenient and adequate test for normality is an examination of the record to find if there are the same number of occurrences on either side of the mean: that is, if the number of positive departures is equal to the number of negative departures. If these numbers are not equal, evidently the distribution is not symmetrical about the mean, and can not be accurately represented by a normal curve. It is not safe to assume that slight deviations from symmetry appearing in short records are due solely to scarcity of observations ("fluctuations of random sampling"), and that the ideal curve will be symmetrical, for in most cases found in actual practice this skewness will persist, regardless of how large the number of observations becomes.

As an example, take the record of annual rainfall at Cumberland, Md., covering a period of 37 years. In 20 of these 37 years the total rainfall was less than the average, and in 17 it was greater than the average.

Now it might seem that if the record were longer, the addition of the other observations would change the position of the mean and the arrangement of the different values sufficiently to overcome this lack of symmetry. But the rainfall record at New Bedford, Mass., covering a period of 94 years, shows that the annual rainfall was less than the average 51 times, and greater than the average only 43 times. From a statistical point of view, a record of even 94 years is too short to furnish more than a rough approximation of the ideal frequencies to be expected, and accordingly the distributions at 21 stations, widely distributed over the United States, with records varying from 30 to 94 years in length, comprising a total of 963 observations of annual rainfall were combined, and even then 581, or 60 per cent, of these 963, were less than the average for their respective stations. Sir Alexander Binnie[9] has assembled the records for 14 stations widely scattered over the world, with periods of observation extending from 19 to 60 years, giving a total of 489 observations, and finds that 265, or 54 per cent of them, were below the average for their stations.

Thus the indication of skewness exhibited in a comparatively short record of 37 years is shown in about the same degree in another record nearly three times as long, and is found to be present not only in records for the entire United States, but for the whole world as well. Although this skewness is so slight that it might easily be ignored in any one record, and in fact some short records do not show this skewness at all, it is evidently not safe to make any deductions based on the assumption that the rainfall at any station for any year is just as likely to be greater as less than the average.

A further difficulty presents itself at this stage, for it is not generally possible to tell exactly, from any a priori considerations, whether the frequency distribution of any variable quantity should follow the normal law. Generally speaking, if an event is the result of a very large number of causes, all these causes independent and each contributing equally toward the resulting event, the distribution will be normal, this being, in fact, the basis of the probability curve and the normal law of error.[10] But if the number of causes is not very large, if they are not all independent of each other, or if some contribute more largely to the result than others, the frequency distribution will exhibit skewness of a greater or less degree. The amount of skewness will depend upon the difference between the nature of these causes and those necessary to give normal distribution. The varying causes of climatic phenomena are not yet known with sufficient exactness to enable one to determine in this manner the character of the distributions which they follow, and since at least 500 to 1,000 observations are necessary to construct a frequency polygon from which an ordinary count of cases would be reliable, and as it has been shown that rainfall, at least, does not follow the normal law, some other means must be devised to secure dependable results.

If we take a number of frequency distributions which tend to be symmetrical and combine them into one distribution, it is evident that the result will still be a symmetrical distribution, and it can be shown that if each of the component distributions is normal the resulting distribution will be normal. Of course, it would be impossible to determine the constants—i. e., the position of

[7] Reed, W. G., & Tolley, H. R. Weather as a business risk in farming. Geogr. Rev., New York, July, 1916. 2: 48–53. Abstract in MONTHLY WEATHER REVIEW, June, 1916 44: 354–355.

[8] Davenport, C. B., op. cit., p. 119; Pearson, K., op. cit., p. 2. For a convenient graphic method of obtaining these ratios see Spillman, W. J., Tolley. H. R., & Reed, W. G. The average interval curve and it applications to meteorological phenomena. MONTHLY WEATHER REVIEW, April, 1916, 44: 197–200.

[9] Binnie, Alexander. Rainfall reservoirs and water supply. New York, 1913. p. 14.
[10] Merriman, M. Textbook on the method of least squares. New York, 1893. p. 15 et seq.

the mean and the size of the standard deviation—for each of the components from the equation of the resulting curve, but the type would still remain. If a number of distributions exhibiting skewness of varying amounts in either direction from the mean are combined, they also will tend to give a normal distribution; but it seems reasonable to suppose that some single type of curve will represent all the distributions of annual rainfall, last killing frost in spring, or any other single phenomenon, and that this type will be that resulting from the combination of a sufficient number of records of varying length selected at random to give a smooth frequency polygon.
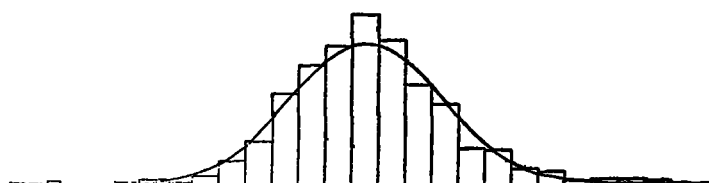

Fig. 1. Frequency polygon and best-fitting normal frequency curve of the date of last killing frost in Spring, for the combined records of 33 stations, comprising 823 observations.

This should be done by combining the departures from the mean for each record, rather than by simply arranging the absolute amounts into a frequency distribution. The means of the records would necessarily have a considerable range, and if the absolute amounts were used it is evident that the distribution resulting from the combination would be much more irregular than if departures were used in each case. If the records are arranged so that all the means coincide, and if the distributions are all of the same type, the only effect of combining them is to smooth out the irregularities in any one of them. Thus this is a kind of artificial method of obtaining an approximation to the general law governing the distribution of a phenomenon when no one record is long enough to show this law conclusively. In the case of last frost in spring the frequency polygon resulting from a combination of the records of 33 stations, with a total of 823 years of observation, was one which followed the normal curve very closely. (See fig. 1.) On account of this fact it was assumed that the distribution for every station would be normal if the number of observations were large enough, and the results of the work based on this assumption seem very satisfactory.[11]
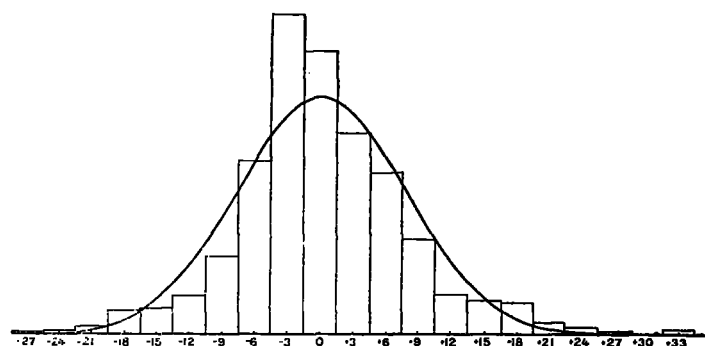

Fig. 2. Frequency polygon (in 3-inch groups) and best-fitting normal frequency curve for the combination of 21 records of annual rainfall, representing 963 observations.

The records of annual rainfall for 21 stations, with periods of observation varying from 30 to 94 years in length, with a total of 963 observations, have been combined in this manner (see fig. 2) to form a frequency polygon, and the best-fitting normal curve found. A comparison of this polygon and the resulting curve with those for the dates of last killing frost in spring (fig. 1)

shows that the normal curve does not describe the fluctuations of annual rainfall nearly so well as it does those of last killing frost in spring. It is possible to determine rigidly the probability that any given distribution would be normal in the limiting case—i. e., the probability that the lack of fit is due merely to paucity of observations. ("The probability that random sampling would lead to as large or larger deviation between theory and observation.")[12] The application of this test shows that it is highly improbable that the normal curve is the limiting form of this polygon.

In a case of this kind, the calculated average has lost two of its most valuable properties. Since the mean, median, and mode are not identical, deviations above and below the mean are not equally likely to occur, and the mean is not the value that will occur most frequently. The only property remaining to the mean is the algebraic one by which it is defined; it is the sum of all the observations divided by their number. In dealing with a phenomenon like rainfall, this property might be put to some use, but it is difficult to see any purpose which it could serve in connection with a phenomenon of temperature. In a like manner, the standard deviation no longer shows the exact way in which the observations are grouped about the mean, and its accuracy as the measure of dispersion becomes less and less as the amount of skewness increases. Looking at it in another way, the position of the mean and the size of the standard deviation no longer define the distribution. In addition, it is necessary to know the position of the median and the mode, and to find some accurate measure of dispersion. The position of the median, that with the same number of observations on each side of it, can be found by counting, but we have no way of determining the probable error of this position, i. e., how much this position would change with an increase in the number of observations, and it is impossible to determine even the approximate position of the mode in a frequency distribution of so few observations as that of the ordinary climatic record.
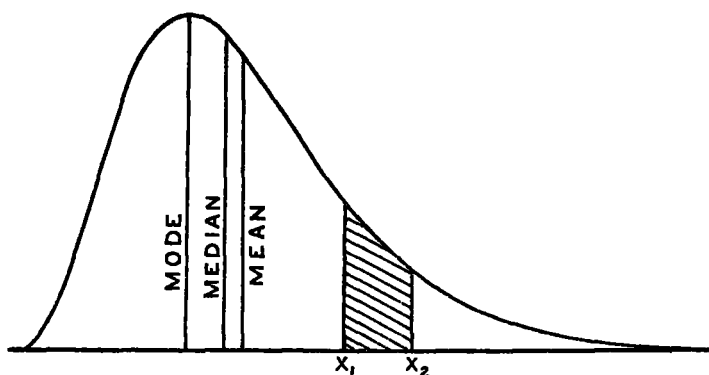
Fig. 3. A typical skew curve. (Compare Marvin, op. cit., p. 555, fig. 3.

If a reliable type of curve can be found to represent all the distributions of a phenomenon, and the constants of its equation determined from each record, the problem will be nearly solved. The position of the mode (see fig. 3) corresponds to the abscissa of the highest ordinate of the curve. The total area under the curve represents the number of observations, and the median is at a point such that its ordinate divides the area into two equal parts. As in the case of the normal curve, the fraction of the area to the right of any ordinate x_1 is the portion of the time that the variable can be expected to have a greater value than x_1, and the area to the left is the portion of the time it can be expected to be less than x_1,

[11] Reed, W. G., & Tolley, H. R. Weather as a business risk in farming. Loc. cit.

[12] Elderton, W. P. Frequency curves and correlation. London, 1906. p. 139 et seq. Pearson, K. Op. cit., 1914, p. 26

while the part between the ordinates of x_1 and x_2 is the portion of the observations which should show a value of the variable between x_1 and x_2. Presumably, all these quantities can be found if the equation of the curve is known.

If this equation be some form, say, $y = F(x)$, then since the position of the mode is at the point where the curve becomes parallel to the axis of x it can be found directly by solving the differential equation $dy/dx = 0$. Any required area can be found by the use of the definite integral; for instance, the area between the ordinates of x_1 and x_2 is $\int_{x_1}^{x_2} F(x)dx$. But the great difficulty lies in obtaining a curve which will give a reasonable fit to the data, and at the same time present an equation that can be differentiated and integrated without such laborious computation as to make the work impracticable.

Various attempts have been made to find usable curves tö describe skew distributions. Pearson [13] has devised a system of curves, of which the normal curve is a special case, designed to cover the entire range of skew variations, and has made extensive use of them in his researches in biology and anthropology. This theory of frequency curves is built around the general differential equation of a unimodal curve. This equation is of the form

$$\frac{1}{y} \cdot \frac{dy}{dx} = \frac{x - a}{F(x)},$$

and if $F(x)$ is expanded by Maclaurin's theorem, it becomes

$$\frac{1}{y} \cdot \frac{dy}{dx} = \frac{x - a}{c_0 + c_1 x + c_2 x^2 + \ldots}$$

The coefficients of the x's can all be determined from the observations, but on account of the uncertainty of the coefficients of the higher powers, only the first three terms of the expansion of $F(x)$ are retained. The differential equation is then integrated to find the general equation of the frequency curve, and the constants of the curve found from the first, second, third, and fourth powers of the departures from the mean of the observations composing the frequency distribution under consideration. However, the resulting equations of the frequency curves are so complicated, and the computation of the constants so lengthy an operation, that the task of analyzing the data at any great number of stations for even one climatic phenomenon would be hopeless. Also it is not possible to integrate the equations of the curves in the general case, and consequently a general expression for the area between any two ordinates can not be obtained. Of course, they give a theoretical position for the median and the mode, and the number of occurrences of any particular amount, but the main part of the problem in dealing with climatic data is to find the portion of the occurrences greater or less than a given amount, and for this they seem to offer little aid.

Others,[14] by an extension of the method from which the normal law was deduced, have developed a theory for skew variation which gives curves more closely allied with the normal curve, and although they do not cover the entire field as thoroughly as do Pearson's formulæ, a

study of the forms of the equations and the calculations of the necessary constants compels one to conclude that in the analysis of climatic data they are much more useful than Pearson's better known forms.

If we call this theory the generalized law of frequency, the equation of the generalized frequency curve is

$$y = A_0 F(x) + A_1 F^{\mathrm{I}}(x) + A_2 F^{\mathrm{II}}(x) + A_3 F^{\mathrm{III}}(x) + A_4 F^{\mathrm{IV}}(x) + \ldots$$

where

$$F(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$$

the equation of the normal curve; $F^{\mathrm{I}}(x)$, $F^{\mathrm{II}}(x)$, etc., are the first, second, and higher derivatives of $F(x)$ and the A's are coefficients independent of x, to be determined from the series of observations under consideration. If the origin is taken at the mean, and if for simplicity, the entire area of the curve is designated as unity, the coefficient A_0 becomes unity, A_1 and A_2 are each equal to zero, and the equation reduces to

$$y = F(x) + A_3 F^{\mathrm{III}}(x) + A_4 F^{\mathrm{IV}}(x) + \ldots$$

The coefficients A_3, A_4, etc., are functions of the third, fourth, and higher powers, respectively, of the departures of the individual observations from their mean.[15] If we call the average of the cubes of the departures μ_3 the average of the fourth powers, μ_4, etc., then

$$A_3 = -\frac{\mu_3}{\underline{|3}}$$

$$A_4 = \frac{\mu_4 - 3\sigma^4}{\underline{|4}}$$

Now if the observations follow the normal law these coefficients, together with all the following ones vanish, and the equation reduces to that of the normal curve. If the distribution is symmetrical, the cubes of all the negative departures will balance the cubes of all the positive departures, and μ_3 will reduce to 0. The value of μ_3 will increase with the skewness, and its sign will show the direction; if the sign is positive, most of the observations will lie below the mean, and consequently the value of the mode will be less than the mean, while if it is negative the reverse will be true. If the skewness is no greater than that found in the distributions of climatic phenomena so far examined, the series is rapidly convergent. The first term, $F(x)$, dominates the entire series, except near the ends of the distribution, and the coefficients of the derivatives of higher order decrease so rapidly in size that the first two or three terms describe the distribution as closely as the number of observations warrant.

The use of the terms of higher order will, of course, give an equation containing a greater number of constants, and on that account will give a closer approximation to any limited set of observations, but the probable error of the constants derived from the higher powers of the departures is large, and because of this it is doubtful if the use of a large number of terms of the series would enable one to make any better estimate of what the frequency distribution of a climatic record would be in the ideal case, than that which can be made from a consideration of only the first (mean), second (standard deviation), and third powers.

[13] *Pearson, K.* Skew variation in homogeneous material. Phil. trans., Roy. Soc., London, 1895, Ser. A, **186**: 343–411.
[14] *Edgeworth, F. Y.* The generalized law of error. Journal, Roy. Stat. Soc., London, 1906, **69**: 497–530.
Thiele, T. N. Theory of observations. London, 1903.
Charlier, C. V. L. Researches into the theory of probability. Lund University. Observatoriet . . . Meddelanden . . . Ser. 2, nr. 4. (Kongl. Fysiografiska Sällskapets Handlingar, N. F., bd. **16**, nr. 5). Lund, 1906.

[15] This method of determining the constants of a curve from the successive powers of the departures is termed the *method of moments* (cf. moments of inertia), for an explanation of which see *Pearson, K.*, The systematic fitting of curves to observations and measurements. Biometrika, v. 1, pp. 265–303; v. 2, pp. 1–23, April Nov., 1902; *Elderton, W. P.*, Frequency curves and correlation. London, 1906. pp. 13–35.

Any presumptive distribution based on the higher coefficients can be of little value as far as these higher coefficients affect it, on account of the fact that the averages of the sums of the higher powers of the departures are liable to great change as the number of observations increase. It has been found that coefficients involving higher powers than the fourth are valueless, even when there are several hundred observations,[16] and consequently the coefficients involving powers higher than the third can be of little value in determining the ideal distribution when the record consists of less than one hundred observations.

The equation then reduces to the form,

$$y = F(x) + A_3 F'''(x)$$

which becomes

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \left[1 - \frac{\mu_3}{|3\sigma^3}\left(\frac{x}{\sigma} - \frac{x^3}{\sigma^3}\right) \right]$$

when the values already defined are given to A_3, $F(x)$, and $F'''(x)$. In order to construct this curve, it is necessary to compute only one constant, μ_3, besides the mean and the standard deviation (see fig. 4).

The equation can be integrated readily and the area between any two ordinates determined. Taking the form

$$y = F(x) + A_3 F''' (x),$$

the indefinite integral is

$$\int y dx = \int F(x)dx + A_3 F'''(x),$$

which upon reduction gives

$$\int_{x_1}^{x_2} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx - \left[\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \cdot \frac{k}{|3}\left(\frac{x^2}{\sigma^2}-1\right) \right]_{x=x_1}^{x=x_2}$$

where $k = \mu_3/\sigma^3$ for the area between the two ordinates whose departures from the mean are (positive) x_1 and x_2. The area beyond (to the right of) the ordinate of any positive departure x_1, is given by

$$\int_{x_1}^{\infty} F(x)dx - \left[\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \cdot \frac{k}{|3}\left(\frac{x^3}{\sigma^2}-1\right) \right]_{x=x_1}^{x=\infty}$$

Since the entire area of the curve has been made equal to unity, the area to the left of the ordinate of x_1 is given by

$$\int_{-\infty}^{x_1} F(x)dx + \left[\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \cdot \frac{k}{|3}\left(\frac{x^3}{\sigma^2}-1\right) \right]_{x=-\infty}^{x=x_1}$$

The sum of the two quantities is the entire area of the curve, and upon addition they give

$$\int_{-\infty}^{x_1} F(x)dx + \int_{x_1}^{\infty} F(x)dx,$$

[16] Pearson, K. Mathematical contributions to the theory of evolution, XIV: The general theory of skew correlation and nonlinear regression. Drapers Co. Research Memoirs, Biometric Series II. London, 1905. p. 7 et seq.
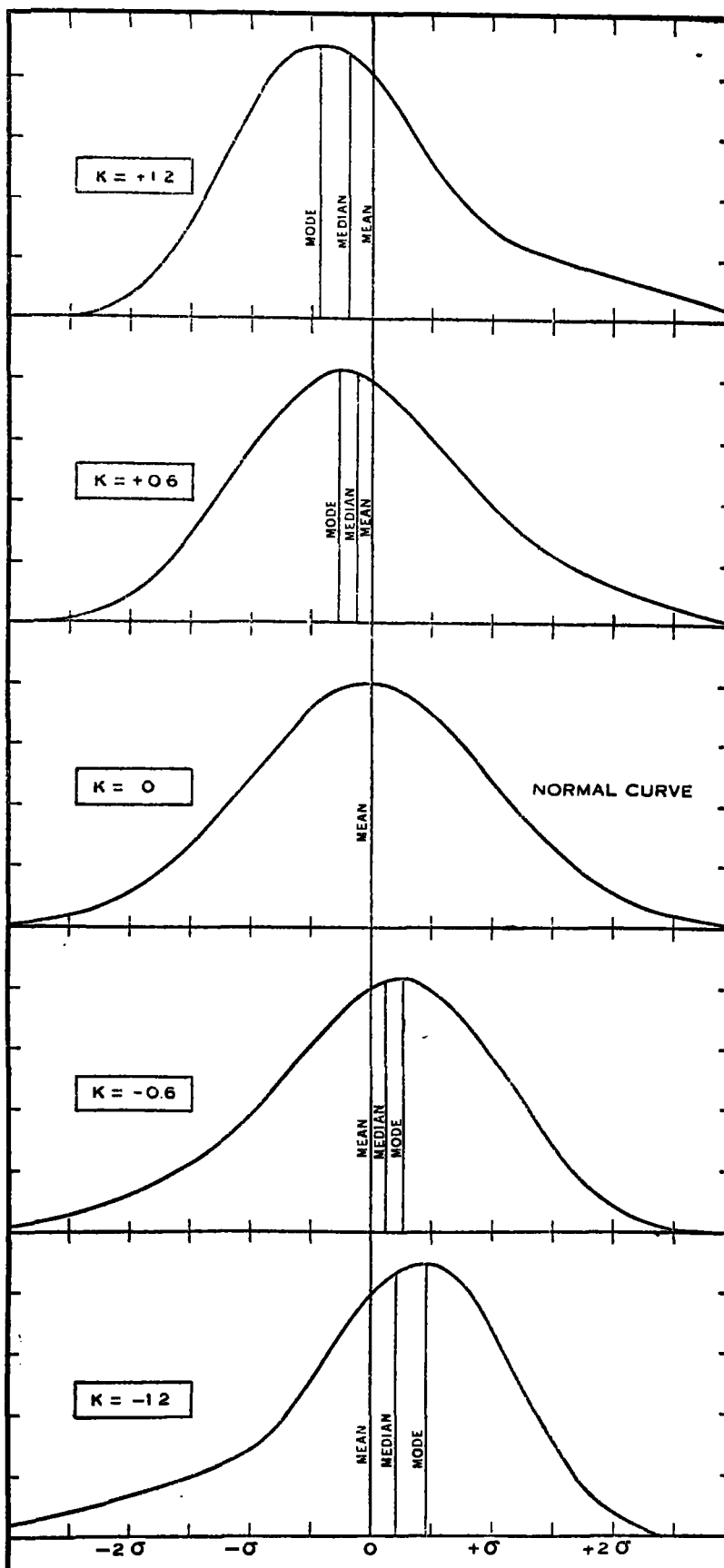

FIG. 4. Frequency curves showing amount and direction of skewness indicated by different values of k.

which is the entire area of the normal curve and equal to unity. The *quantity enclosed in the brackets* vanishes when $x = \infty$, and it can be evaluated by direct computation [17] for different values of k and σ. The first term $\int F(x)dx$ is the integral of the equation of the normal curve, and its value can be obtained from the tables of Pearson or Davenport. The two terms can then be combined to form new tables, which will give immediately the portion of the area of the skew curve on either side of any ordinate when k and σ are known for the distribution. Tables giving the percentage of the area of the curve to the left of the ordinate whose abscissa is x, have been prepared for values of k, at intervals of 0.2 from $k = -1.4$ to $k = +1.4$, and values of x/σ at intervals of 0.05, from $x/\sigma = -3.00$ to $x/\sigma = +3.00$. (Table 1.) By interpolation from this table, it is possible to find, to four places of decimals, the portion of the area on either side of any ordinate or that between any two ordinates, or what is the same thing, the probability of an observation being greater or less than any given value, or the probability of its lying between any two selected values.

The reciprocal of the probability of the occurrence of an event may be called the *average interval* between such occurrences. For instance, if it has been found that one-fifth of the observations of annual rainfall at a station should be below a certain amount, then the probability of the rainfall for any year being less than this amount is one-fifth, and the average interval between occurrences of rainfall less than this amount is five years. If, then, the reciprocals of the quantities (considered as percentages) given in Table 1 are plotted against their respective values of x/σ, the resulting family of curves could be used to find the average interval between occurrences greater or less than any selected amount.[18] Such curves are easily constructed and in many cases where meteorological data are under consideration, their use will be found preferable to that of the tables. The use of these curves, to find the appropriate value of x/σ for any given average interval, or vice versa, when k is known, requires less time than the use of the tables, and if the curves are properly constructed, values thus found will generally have an accuracy comparable to that of the observations themselves.

The use of the skew curves, and the resulting table, has been tested in the case of 38 records of minimum winter temperatures, in an attempt to find the values below which the temperature should not be expected to fall oftener than once in 10 years on the average. There is a lack of symmetry in the frequency distributions of winter minima about equal to that in those of annual rainfall, but it is in the opposite direction, the greater number of the observations being above the mean instead of below it. The departure, below the mean, of the value which the winter minimum will exceed in one-tenth of the years, or below which the probability of an occurrence is 1/10, will be represented by the abscissa whose ordinate cuts off 10% of the area from the lower (left) end of the frequency curve derived from the observations. The actual computations in any case necessary to find this value would be somewhat as follows (see Table 2):

(1) Find the mean, M_0, the standard deviation, σ, and the average of the cubes of the departures from the mean, μ_3. In the example, $M_0 = +16.67°F.$, $\sigma = 7.37$, and $\mu_3 = -276.23$.

(2) Find the value of k, from the formula $k = \mu_3/\sigma^3$. In the example, $k = -0.69$.

(3) For the value of k thus found, determine from Table 1 the value of x/σ which corresponds to 10% of the area of the curve. For $k = -0.69$, this value is -1.378.

(4) Multiply the standard deviation, σ, by the value of x/σ thus found and subtract the product from the mean, M_0. For Portland, Oreg., the resulting value is $+6.51°F.$ The probability that the temperature will fall below the value thus obtained is 1/10, or what is the same thing, the temperature should fall below this value once in 10 years on the average.

An abstract of the computation of this value for each of the 38 selected stations is given in Table 3. An examination of the k's for the different records bears out the statement already made that there are generally a greater number of observations above the mean than below it. At all except four of the stations, k is negative. The sign of k depends upon the sign of μ_3, and as already stated, if this sign is negative, most of the observations lie above the mean and the value of the mode is greater than that of the mean.

These computed values can not be expected to agree exactly with those found by a simple count of the lowest observed values in each record, i. e., although all the records examined were approximately 40 years in length, it is not to be expected that the computed values below which the temperature should fall in 1/10 of the years would have been exceeded exactly four times in each record. However, the differences between these computed values and those found by counting should be promiscuous, and the two methods should agree in the aggregate. If we add together the number of observations in each record which fall below this computed value the sum should be very nearly one-tenth of the total number of observations. In these 38 records, with a total of 1,519 observations, 150 observations fell below the values computed by this method. Similar values were computed at each station, on the assumption that the distribution was normal ($k = 0$), by simply subtracting 1.28σ from the mean, and it was found that 173 of the 1,519 observations fell below these values. This shows that the skew curves are much more accurate, at least for this purpose, than the normal curve.

The method was tested further by computing, for each of these stations, the value below which the temperature should fall 9 years in 10 on the average. The probability that the minimum temperature will be above such a value is 1/10, and the departure above the mean of that value will be represented by the abscissa whose ordinate cuts off 10% of the area from the upper (right) end of the frequency curve. The procedure then was the same as that outlined above, except that the value of x/σ which corresponds to 90% of the area of the curve was determined for the proper value of k in each case, and the resulting value of x was added to the mean instead of subtracted from it. Here it was found that 144 of the 1,519 observations were above the values thus computed. This is not quite such close agreement as that found above, but it is considerably better than that given by assuming normal distribution. There were only 127 occurrences above the values computed by adding 1.28σ to the mean in each case. Thus, in an endeavor to find values that should be exceeded in 10 per cent of the cases, the use of the normal curve gave values which were exceeded by only 8.3 per cent of the occurrences; while the use of the skew curves, where k as well as σ was taken into account, gave values that were exceeded by 9.5 per cent of the occurrences.

---

[17] For a table of values of the quantity $\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{\sigma^2}}$, for varying values of $\frac{x}{\sigma}$, see Pearson, K. Tables for Statisticians and Biometricians. Cambridge, 1914. Table II, pp. 2-8.

[18] For an explanation of the construction and some of the uses of the average interval curve for normal distributions see *Spillman, W. J., Tolley, H. R., & Reed, W. G.*, The average interval curve and its application to meteorological phenomena, MONTHLY WEATHER REVIEW April, 1916, 44: 197-200.

TABLE 1.—*Area of skew frequency curves in terms of the abscissa.*

(Total area of curve=100.)

$\frac{x}{\sigma}$	−1.4	−1.2	−1.0	−0.8	−0.6	−0.4	−0.2	0.0	+0.2	+0.4	+0.6	+0.8	+1.0	+1.2	+1.4
−3.00	0.962	0.844	0.726	0.608	0.490	0.371	0.253	0.135	0.017						
−2.95	1.083	.951	.819	.687	.555	.423	.291	.159	.027						
−2.90	1.216	1.069	.922	.775	.628	.481	.334	.187	.040						
−2.85	1.361	1.198	1.034	.871	.708	.545	.382	.219	.055						
−2.80	1.519	1.338	1.158	.977	.797	.616	.436	.256	.075						
−2.75	1.690	1.492	1.293	1.094	0.895	0.696	0.497	0.298	0.099						
−2.70	1.876	1.658	1.439	1.221	1.002	.784	.565	.347	.128						
−2.65	2.076	1.837	1.598	1.359	1.120	.881	.642	.402	.163						
−2.60	2.292	2.031	1.770	1.509	1.248	.988	.727	.466	.205						
−2.55	2.522	2.239	1.955	1.672	1.389	1.105	.822	.539	.255						
−2.50	2.768	2.461	2.155	1.848	1.541	1.234	0.928	0.621	0.314	0.007					
−2.45	3.030	2.699	2.368	2.037	1.707	1.376	1.045	.714	.383	.053					
−2.40	3.307	2.952	2.596	2.241	1.886	1.530	1.175	.820	.464	.109					
−2.35	3.600	3.220	2.839	2.459	2.079	1.699	1.319	.939	.559	.178					
−2.30	3.908	3.503	3.098	2.693	2.288	1.883	1.477	1.072	.667	.262					
−2.25	4.231	3.801	3.371	2.942	2.512	2.082	1.652	1.222	0.793	0.363					
−2.20	4.569	4.115	3.661	3.207	2.753	2.298	1.844	1.390	.936	.482	.028				
−2.15	4.921	4.445	3.966	3.488	3.010	2.533	2.055	1.578	1.100	.623	.145				
−2.10	5.286	4.786	4.286	3.786	3.286	2.786	2.286	1.786	1.286	.787	.287				
−2.05	5.664	5.143	4.622	4.102	3.581	3.060	2.539	2.018	1.497	.977	.456				
−2.00	6.054	5.514	4.975	4.435	3.895	3.355	2.815	2.275	1.735	1.195	0.655	0.115			
−1.95	6.456	5.899	5.342	4.786	4.229	3.672	3.116	2.559	2.002	1.445	.889	.332			
−1.90	6.868	6.297	5.726	5.155	4.584	4.013	3.443	2.872	2.301	1.730	1.159	.588	0.017		
−1.85	7.289	6.707	6.125	5.543	4.961	4.380	3.798	3.216	2.634	2.052	1.470	.888	.306		
−1.80	7.720	7.130	6.540	5.951	5.362	4.772	4.183	3.593	3.004	2.414	1.825	1.235	.646	0.056	
−1.75	8.158	7.565	6.972	6.379	5.785	5.192	4.599	4.006	3.413	2.820	2.226	1.633	1.040	0.447	
−1.70	8.604	8.012	7.419	6.827	6.234	5.642	5.049	4.457	3.864	3.272	2.679	2.087	1.494	.901	0.309
−1.65	9.057	8.470	7.883	7.296	6.709	6.121	5.534	4.947	4.360	3.773	3.186	2.598	2.011	1.424	.837
−1.60	9.517	8.941	8.364	7.787	7.210	6.634	6.057	5.480	4.903	4.326	3.750	3.173	2.596	2.019	1.442
−1.55	9.984	9.423	8.862	8.301	7.740	7.179	6.618	6.057	5.496	4.935	4.374	3.813	3.252	2.691	2.130
−1.50	10.458	9.919	9.379	8.839	8.300	7.760	7.220	6.681	6.141	5.601	5.062	4.522	3.982	3.443	2.903
−1.45	10.940	10.427	9.915	9.403	8.890	8.378	7.865	7.353	6.841	6.328	5.816	5.303	4.791	4.278	3.766
−1.40	11.430	10.950	10.471	9.992	9.513	9.034	8.555	8.076	7.597	7.117	6.638	6.159	5.680	5.201	4.722
−1.35	11.929	11.489	11.049	10.610	10.170	9.730	9.291	8.851	8.411	7.971	7.532	7.092	6.652	6.212	5.773
−1.30	12.439	12.045	11.651	11.257	10.862	10.468	10.074	9.680	9.286	8.892	8.498	8.103	7.709	7.315	6.921
−1.25	12.962	12.620	12.277	11.935	11.592	11.250	10.907	10.565	10.223	9.880	9.538	9.195	8.853	8.510	8.168
−1.20	13.501	13.216	12.931	12.646	12.361	12.077	11.792	11.507	11.222	10.937	10.653	10.368	10.083	9.798	9.513
−1.15	14.057	13.835	13.614	13.393	13.171	12.950	12.729	12.507	12.286	12.064	11.843	11.622	11.400	11.179	10.958
−1.10	14.634	14.482	14.329	14.177	14.024	13.872	13.719	13.567	13.414	13.262	13.109	12.957	12.804	12.652	12.499
−1.05	15.236	15.157	15.079	15.000	14.922	14.843	14.764	14.686	14.607	14.529	14.450	14.372	14.293	14.215	14.136
−1.00	15.866	15.866	15.866	15.866	15.866	15.866	15.866	15.866	15.866	15.866	15.866	15.866	15.866	15.866	15.866
−.95	16.528	16.610	16.693	16.775	16.858	16.940	17.023	17.106	17.188	17.271	17.353	17.436	17.518	17.601	17.684
−.90	17.226	17.395	17.563	17.732	17.900	18.069	18.237	18.406	18.575	18.743	18.912	19.080	19.249	19.417	19.586
−.85	17.946	18.223	18.481	18.738	18.995	19.252	19.509	19.766	20.023	20.281	20.548	20.795	21.052	21.309	21.566
−.80	18.752	19.100	19.447	19.795	20.143	20.490	20.838	21.186	21.533	21.881	22.228	22.576	22.924	23.271	23.619
−0.75	19.589	20.028	20.467	20.906	21.345	21.784	22.224	22.663	23.102	23.541	23.980	24.419	24.859	25.298	25.737
−.70	20.481	21.011	21.542	22.073	22.604	23.135	23.666	24.196	24.727	25.258	25.789	26.320	26.851	27.381	27.912
−.65	21.433	22.054	22.676	23.298	23.919	24.541	25.163	25.785	26.406	27.028	27.650	28.272	28.893	29.515	30.137
−.60	22.449	23.160	23.871	24.582	25.293	26.004	26.714	27.425	28.136	28.847	29.558	30.269	30.980	31.691	32.401
−.55	23.535	24.332	25.129	25.927	26.724	27.521	28.319	29.116	29.913	30.711	31.508	32.305	33.103	33.900	34.697
−0.50	24.693	25.573	26.453	27.333	28.213	29.093	29.974	30.854	31.734	32.614	33.494	34.374	35.255	36.135	37.015
−.45	25.927	26.885	27.844	28.802	29.760	30.719	31.677	32.636	33.594	34.552	35.511	36.469	37.428	38.386	39.344
−.40	27.240	28.271	29.302	30.333	31.364	32.396	33.427	34.458	35.489	36.520	37.551	38.582	39.614	40.645	41.676
−.35	28.634	29.731	30.829	31.927	33.024	34.122	35.219	36.317	37.415	38.512	39.610	40.707	41.805	42.902	44.000
−.30	30.111	31.268	32.424	33.581	34.738	35.895	37.052	38.209	39.366	40.523	41.679	42.836	43.993	45.150	46.307
−0.25	31.671	32.879	34.088	35.296	36.504	37.713	38.921	40.129	41.338	42.546	43.754	44.963	46.171	47.379	48.588
−.20	33.315	34.566	35.817	37.069	38.320	39.571	40.823	42.074	43.325	44.577	45.828	47.079	48.331	49.582	50.833
−.15	35.041	36.326	37.612	38.897	40.182	41.468	42.753	44.038	45.324	46.609	47.894	49.180	50.465	51.750	53.036
−.10	36.848	38.158	39.468	40.777	42.087	43.397	44.707	46.017	47.327	48.637	49.947	51.257	52.567	53.877	55.187
−.05	38.732	40.057	41.382	42.707	44.032	45.356	46.681	48.006	49.331	50.656	51.981	53.305	54.630	55.955	57.280
0.00	40.691	42.021	43.351	44.681	46.011	47.340	48.670	50.000	51.330	52.660	53.989	55.319	56.649	57.979	59.3__

TABLE 1.—*Area of skew frequency curves in terms of the abscissa*—Continued.

(Total area of curve=100.)

Values of k.

$\frac{x}{\sigma}$	−1.4	−1.2	−1.0	−0.8	−0.6	−0.4	−0.2	−0.0	+0.2	+0.4	+0.6	+0.8	+1.0	+1.2	+1.4
0.00	40.691	42.021	43.251	44.681	46.011	47.340	48.670	50.000	51.330	52.660	53.989	55.319	56.649	57.979	59.309
+0.05	42.720	44.045	45.370	46.695	48.019	49.344	50.669	51.994	53.319	54.644	55.968	57.293	58.618	59.943	61.268
.10	44.813	46.123	47.473	48.743	50.053	51.363	52.673	53.983	55.293	56.603	57.913	59.223	60.532	61.812	63.152
.15	46.964	48.250	49.535	50.820	52.106	53.391	54.676	55.982	57.247	58.532	59.818	61.103	62.388	63.674	64.959
.20	49.167	50.418	51.669	52.921	54.172	55.423	56.675	57.926	59.177	60.429	61.680	62.931	64.183	65.434	66.685
.25	51.412	52.621	53.829	55.037	56.246	57.454	58.662	59.871	61.079	62.287	63.496	64.704	65.912	67.121	68.329
+0.30	53.693	54.850	56.007	57.164	58.321	59.477	60.634	61.791	62.948	64.105	65.262	66.419	67.576	68.732	69.889
.35	56.000	57.098	58.195	59.293	60.390	61.488	62.585	63.683	64.781	65.878	66.976	68.073	69.171	70.269	71.366
.40	58.324	59.355	60.386	61.418	62.449	63.480	64.511	65.542	66.573	67.604	68.636	69.667	70.698	71.729	72.760
.45	60.656	61.614	62.572	63.531	64.489	65.448	66.406	67.364	68.323	69.281	70.240	71.198	72.156	73.115	74.073
.50	62.985	63.865	64.745	65.626	66.506	67.386	68.266	69.146	70.026	70.907	71.787	72.667	73.547	74.427	75.307
+0.55	65.303	66.100	66.897	67.695	68.492	69.289	70.087	70.884	71.681	72.479	73.276	74.073	74.871	75.668	76.465
.60	67.599	68.309	69.020	69.731	70.442	71.153	71.864	72.575	73.286	73.996	74.707	75.418	76.129	76.840	77.551
.65	69.863	70.485	71.107	71.728	72.350	72.972	73.594	74.215	74.837	75.459	76.081	76.702	77.324	77.946	78.567
.70	72.088	72.619	73.149	73.680	74.211	74.742	75.273	75.804	76.334	76.865	77.396	77.927	78.458	78.989	79.519
.75	74.263	74.702	75.141	75.581	76.020	76.459	76.898	77.337	77.776	78.216	78.655	79.094	79.533	79.972	80.411
+0.80	76.381	76.729	77.076	77.424	77.772	78.119	78.467	78.814	79.162	79.510	79.857	80.205	80.553	80.900	81.248
.85	78.434	78.691	78.948	79.205	79.462	79.719	79.977	80.234	80.491	80.748	81.005	81.262	81.519	81.777	82.034
.90	80.411	80.585	80.751	80.920	81.088	81.257	81.425	81.594	81.763	81.931	82.100	82.268	82.437	82.605	82.774
.95	82.316	82.399	82.481	82.564	82.617	82.729	82.812	82.894	82.977	83.060	83.142	83.225	83.307	83.390	83.472
1.00	84.134	84.134	84.134	84.134	84.134	84.134	84.134	84.134	84.134	84.134	84.134	84.134	84.134	84.134	84.134
+1.05	85.834	85.785	85.707	85.629	85.550	85.471	85.393	85.314	85.236	85.157	85.078	85.000	84.921	84.843	84.764
1.10	87.501	87.345	87.196	87.043	86.891	86.738	86.586	86.433	86.281	86.128	85.976	85.823	85.671	85.518	85.366
1.15	89.042	88.821	88.600	88.378	88.157	87.935	87.714	87.493	87.271	87.050	86.829	86.607	86.386	86.165	85.943
1.20	90.487	90.202	89.917	89.632	89.347	89.063	88.778	88.493	88.208	87.923	87.639	87.354	87.069	86.784	86.499
1.25	91.832	91.490	91.147	90.805	90.462	90.120	89.777	89.435	89.093	88.750	88.408	88.065	87.723	87.380	87.038
+1.30	93.079	92.685	92.291	91.897	91.502	91.108	90.714	90.320	89.926	89.532	89.138	88.743	88.349	87.955	87.561
1.35	94.227	93.788	93.348	92.908	92.468	92.029	91.589	91.149	90.709	90.270	89.830	89.390	88.951	88.511	88.071
1.40	95.278	94.799	94.320	93.841	93.362	92.883	92.404	91.924	91.445	90.966	90.487	90.008	89.529	89.050	88.570
1.45	96.234	95.722	95.209	94.697	94.184	93.672	93.159	92.647	92.135	91.622	91.110	90.597	90.085	89.573	89.060
1.50	97.097	96.557	96.018	95.478	94.938	94.399	93.859	93.319	92.780	92.240	91.700	91.161	90.621	90.081	89.542
+1.55	97.870	97.309	96.748	96.187	95.626	95.065	94.504	93.943	93.382	92.821	92.260	91.699	91.138	90.577	90.016
1.60	98.558	97.981	97.404	96.827	96.250	95.674	95.097	94.520	93.943	93.366	92.790	92.213	91.636	91.059	90.483
1.65	99.163	98.576	97.989	97.402	96.814	96.227	95.640	95.053	94.466	93.879	93.291	92.704	92.117	91.530	90.943
1.70	99.691	99.099	98.506	97.913	97.321	96.728	96.136	95.543	94.951	94.358	93.766	93.173	92.581	91.988	91.396
1.75		99.553	98.960	98.367	97.774	97.180	96.587	95.994	95.401	94.808	94.215	93.621	93.028	92.435	91.842
+1.80		99.944	99.354	98.765	98.175	97.586	96.996	96.407	95.817	95.228	94.638	94.049	93.460	92.870	92.281
1.85			99.694	99.112	98.530	97.948	97.366	96.784	96.202	95.620	95.039	94.457	93.875	93.293	92.711
1.90			99.983	99.412	98.841	98.270	97.699	97.128	96.557	95.987	95.416	94.845	94.274	93.703	93.132
1.95				99.668	99.111	98.555	97.998	97.441	96.884	96.328	95.771	95.214	94.658	94.101	93.544
2.00				99.885	99.345	98.805	98.265	97.725	97.185	96.645	96.105	95.565	95.025	94.486	93.946
+2.05					99.544	99.023	98.503	97.982	97.461	96.940	96.419	95.898	95.378	94.857	94.336
2.10					99.713	99.213	98.714	98.214	97.714	97.214	96.714	96.214	95.714	95.214	94.714
2.15					99.855	99.377	98.900	98.422	97.945	97.467	96.990	96.512	96.034	95.557	95.079
2.20					99.972	99.518	99.064	98.610	98.156	97.702	97.247	96.793	96.339	95.885	95.431
2.25						99.637	99.207	98.778	98.348	97.918	97.488	97.058	96.629	96.199	95.769
+2.30						99.738	99.333	98.928	98.523	98.117	97.712	97.307	96.902	96.497	96.092
2.35						99.822	99.441	99.061	98.681	98.301	97.921	97.541	97.161	96.780	96.400
2.40						99.891	99.536	99.180	98.825	98.470	98.114	97.759	97.404	97.048	96.693
2.45						99.947	99.617	99.286	98.955	98.624	98.293	97.963	97.632	97.301	96.970
2.50						99.993	99.686	99.379	99.072	98.766	98.459	98.152	97.845	97.539	97.232
+2.55							99.745	99.461	99.178	98.895	98.611	98.328	98.045	97.761	97.478
2.60							99.795	99.534	99.273	99.012	98.752	98.491	98.230	97.969	97.708
2.65							99.837	99.598	99.358	99.119	98.880	98.641	98.402	98.163	97.924
2.70							99.872	99.653	99.435	99.216	98.998	98.779	98.561	98.342	98.124
2.75							99.901	99.702	99.503	99.304	99.105	98.906	98.707	98.508	98.310
+2.80							99.925	99.744	99.564	99.383	99.203	99.023	98.842	98.662	98.481
2.85							99.945	99.781	99.618	99.455	99.292	99.129	98.966	98.802	98.639
2.90							99.960	99.813	99.666	99.519	99.372	99.225	99.078	98.931	98.784
2.95							99.973	99.841	99.709	99.577	99.445	99.313	99.181	99.049	98.917
3.00							99.983	99.865	99.747	99.629	99.510	99.392	99.274	99.156	99.038

TABLE 2.—*Calculation of value below which the winter minimum will fall once in 10 years, on the average, at Portland, Oreg.*

Year.	Minimum temperature.	d.	d².	d³.
	°F.	°F.		
1888	-2	-19	361	-6,859
1875	3	-14	196	-2,744
1879	3	-14	196	-2,744
1909	6	-11	121	-1,331
1883	7	-10	100	-1,000
1884	7	-10	100	-1,000
1893	8	-9	81	-729
1887	9	-8	64	-512
1899	9	-8	64	-512
1890	10	-7	49	-343
1896	11	-6	36	-216
1902	13	-4	16	-64
1907	13	-4	16	-64
1886	15	-2	4	-8
1885	17	0	0	0
1905	17	0	0	0
1878	18	1	1	1
1882	18	1	1	1
1894	18	1	1	1
1880	19	2	4	8
1900	19	2	4	8
1876	20	3	9	27
1892	20	3	9	27
1912	20	3	9	27
1898	21	4	16	64
1910	21	4	16	64
1897	22	5	25	125
1906	22	5	25	125
1913	22	5	25	125
1889	23	6	36	216
1891	23	6	36	216
1908	23	6	36	216
1911	23	6	36	216
1881	24	7	49	343
1903	24	7	49	343
1877	25	8	64	512
1895	25	8	64	512
1901	26	9	81	729
1904	28	11	121	1,331
Sums		-13	2,121	-12,889
Mean	16.67			

COMPUTATION.

Quantity.	Symbol.	Value.
Number of years of observation	n	39
Mean winter minimum	M_0	+16.67°F.
Convenient number near the mean	M	+17°
Departure from M	d	
Sum of column d	Σd	-13
Average departure from M	$\Sigma d/n$	-0.33
Sum of column d^2	Σd^2	+2,121
Average of square of departures from M	$\Sigma d^2/n$	+54.38
Average of square of departures from the mean [1]	$\dfrac{\Sigma d^2}{n}-\left(\dfrac{\Sigma d}{n}\right)^2$	+54.27
Standard deviation	$\sigma=\sqrt{\dfrac{\Sigma d^2}{n}-\left(\dfrac{\Sigma d}{n}\right)^2}$	7.37
Sum of column d^3	Σd^3	-12,889
Average of cube of departures from M	$\Sigma d^3/n$	-330.48
Average of cube of departures from the mean [1]	$\mu_3=\dfrac{\Sigma d^3}{n}-3\left(\dfrac{\Sigma d}{n}\right)\left(\dfrac{\Sigma d^2}{n}\right)+2\left(\dfrac{\Sigma d}{n}\right)^3$	-276.23
k	$k=\mu_3/\sigma^3$	-0.69
Value from Table 1, for $k=-0.69$, and *average interval*=10	x/σ	-1.378
Departure below mean that will be exceeded in 1/10 of the years	$x=1.378\sigma$	10.16°F.
Value below which winter minimum will fall in 1/10 of the years	M_0-x	+6.51°F.

[1] See *Davenport, C. B.*: Statistical methods, ed. 3, 1914, pp. 20–21, for formulæ reducing these quantities to the true mean. The notation in the table is different from that used by Davenport.

TABLE 3.—*Abstract of computation of values of minimum winter temperatures (t) which should be exceeded, on the average, once in 10 years.*

State.	Station.	Number of observations, n.	Mean minimum temperature.	σ	k	$\dfrac{x}{\sigma}$ (from Table 1).	Departures of t from mean.	t.	Number of observations below t.
			°F.				°F.	°F.	
Alabama	Mobile	42	+21.19	6.29	-0.93	1.427	-8.98	+12.21	3
Do	Montgomery	41	+16.95	6.30	-1.02	1.448	-9.12	+7.83	2
California	San Diego	42	+35.62	2.95	-0.45	1.338	-3.95	+31.67	1
Do	San Francisco	39	+37.00	2.74	-0.33	1.322	-3.62	+33.38	4
Dist. of Columbia.	Washington	39	+2.85	7.33	-0.85	1.411	-10.34	-7.49	4
Florida	Key West	39	+50.64	3.89	-0.41	1.332	-5.18	+45.46	5
Georgia	Augusta	39	+16.97	5.45	-0.66	1.373	-7.48	+9.49	5
Do	Savannah	39	+21.51	5.29	-0.65	1.371	-7.25	+14.26	4
Illinois	Cairo	42	+0.26	7.09	-0.23	1.308	-9.27	-9.01	3
Indiana	Indianapolis	42	-8.88	6.63	-0.13	1.296	-8.58	-17.46	4
Iowa	Davenport	41	-15.02	6.53	0.00	1.282	-8.36	-23.38	4
Do	Des Moines	35	-17.91	6.30	-0.21	1.306	-8.23	-26.14	4
Do	Keokuk	42	-13.36	6.73	-0.17	1.302	-8.76	-22.12	4
Kansas	Dodge City	39	-10.67	7.01	+0.02	1.279	-8.97	-19.64	3
Kentucky	Louisville	41	-0.88	7.29	-0.39	1.330	-9.70	-10.58	2
Louisiana	New Orleans	39	+25.80	5.72	-0.99	1.440	-8.24	+17.56	3
Do	Shreveport	39	+15.38	6.97	-0.67	1.375	-9.58	+5.80	3
Maryland	Baltimore	43	+5.63	5.63	-0.76	1.390	-7.83	-2.20	4
Mississippi	Vicksburg	39	+16.90	6.36	-0.86	1.411	-8.97	+7.93	3
Nebraska	North Platte	39	-18.69	8.35	-0.12	1.295	-10.81	-29.50	3
Do	Omaha	39	-16.10	6.77	-0.32	1.320	-8.94	-25.04	4
New Jersey	Atlantic City	39	+3.62	5.58	+0.07	1.274	-7.11	-3.49	5
New Mexico	Santa Fe	39	-2.46	5.56	-0.15	1.298	-7.22	-9.68	5
New York	New York	42	+1.95	4.79	+0.27	1.256	-6.02	-4.07	4
North Carolina	Wilmington	39	+17.08	4.82	-0.17	1.301	-6.27	+10.81	5
Ohio	Cincinnati	43	-2.28	6.25	-0.18	1.302	-8.14	-10.42	4
Oregon	Portland	39	+16.67	7.37	-0.69	1.378	-10.16	+6.51	4
Pennsylvania	Pittsburgh	39	-2.69	5.45	-0.68	1.377	-7.50	-10.19	3
South Carolina	Charleston	39	+21.69	5.50	-0.81	1.380	-7.59	+14.10	6
Tennessee	Knoxville	43	+4.14	8.43	-0.59	1.362	-11.48	-7.34	5
Do	Memphis	39	+7.97	7.00	-0.36	1.327	-9.29	-1.32	4
Do	Nashville	39	+3.33	7.73	-0.15	1.299	-10.04	-6.71	4
Texas	Galveston	42	+25.33	6.36	-0.44	1.337	-8.50	+16.83	3
Utah	Salt Lake City	39	+0.05	6.72	-1.16	1.459	-9.82	-9.77	5
Virginia	Cape Henry	39	+14.59	4.74	-0.23	1.309	-6.20	+8.39	5
Do	Lynchburg	39	+6.10	6.68	-0.23	1.308	-8.74	-2.64	7
Do	Norfolk	39	+13.64	5.12	-0.30	1.318	-6.75	+6.89	5
Wyoming	Cheyenne	41	-18.71	7.78	-0.11	1.294	-10.07	-28.78	4

GRAPHIC METHOD OF REPRESENTING AND COMPARING DROUGHT INTENSITIES.[1]

By THORNTON T. MUNGER.

[U. S. Forest Service, Portland, Oreg., Nov. 1, 1915.]

It is a matter of interest among foresters to find a way for expressing in some graphic quantitative fashion the comparative forest fire risk of various years, and to determine the relative fire risk in various regions. There are so many factors that combine to create a fire hazard in our forests that it is difficult to express them in a statistical or graphic form.

The most influential meteorological factors are the infrequency of soaking rains, the total amount of rain in the dry season, the depth of the winter snow and the time of its disappearance, the humidity of the atmosphere, the frequency of very hot days, the occurrence of high winds, particularly of dry winds, and the seasonal temperatures as they affect the time at which the herbaceous vegetation matures and dries up. All these factors of precipitation, temperature, and wind movement are so complexly interwoven that it seems to be impossible to combine them and consider them jointly. The one single factor that has the most important influence on

[1] This method of showing drought severity was described by District Forecaster E. A. Beals at the meeting of the Western Forestry and Conservation Association on Dec. 7, 1914, using diagrams modeled after those originated by the author.